

**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>7</sup> : <b>C12N 15/29, 15/61, 15/82, 5/10, C07K 14/415, A01H 5/00</b>		A1	(11) International Publication Number: <b>WO 00/11177</b> (43) International Publication Date: <b>2 March 2000 (02.03.00)</b>
(21) International Application Number: <b>PCT/US99/18628</b>		(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).	
(22) International Filing Date: <b>17 August 1999 (17.08.99)</b>			
(30) Priority Data: <b>60/097,233 20 August 1998 (20.08.98) US</b>			
(71) Applicant: <b>PIONEER HI-BRED INTERNATIONAL, INC.</b> [US/US]; 800 Capital Square, 400 Locust Street, Des Moines, IA 50309 (US).			
(72) Inventors: <b>MARTINO-CATT, Susan, J.; 499 N.W. 69th Place, Ankeny, IA 50021 (US). LAPPEGARD, Kathryn, K.; 1228 4th Street, Nevada, IA 50201 (US). WANG, XUN; 12524 Caminito Vista Soledad, San Diego, CA 92130 (US). BOWEN, Benjamin, A.; 7027 Buckingham Boulevard, Berkeley, CA 94705 (US).</b>		Published <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>	
(74) Agents: <b>MICHEL, Marianne, H. et al.; Darwin Building, 7100 N.W. 62nd Avenue, Johnston, IA 50131-1000 (US).</b>			
(54) Title: <b>SEED-PREFERRED PROMOTERS</b>			
(57) Abstract			
<p>The present invention provides compositions and methods for regulating expression of heterologous nucleotide sequences in a plant. Compositions are novel nucleotide sequences for seed-preferred promoters isolated from genes for Cim1 (cytokinin-induced message); cZ19B1 (maize 19kDa zein); or milps (myoinositol-1-phosphate synthase). A method for expressing a heterologous nucleotide sequence in a plant using the promoter sequences disclosed herein is provided. The method comprises transforming a plant cell to comprise a heterologous nucleotide sequence operably linked to one of the seed-preferred promoters of the present invention and regenerating a stably transformed plant from the transformed plant cell.</p>			

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		

## SEED-PREFERRED PROMOTERS

## FIELD OF THE INVENTION

The present invention relates to the field of plant molecular biology, more  
5 particularly to regulation of gene expression in plants.

## BACKGROUND OF THE INVENTION

Expression of heterologous DNA sequences in a plant host is dependent upon the presence of an operably linked promoter that is functional within the plant  
10 host. Choice of the promoter sequence will determine when and where within the organism the heterologous DNA sequence is expressed. Where continuous expression is desired throughout the cells of a plant, constitutive promoters are utilized. In contrast, where gene expression in response to a stimulus is desired, inducible promoters are the regulatory element of choice. Where expression in  
15 specific tissues or organs are desired, tissue-specific promoters may be used. That is, they may drive expression in specific tissues or organs. Such tissue-specific promoters may be constitutive or inducible. In either case, additional regulatory sequences upstream and/or downstream from the core promoter sequence may be included in expression constructs of transformation vectors to  
20 bring about varying levels of expression of heterologous nucleotide sequences in a transgenic plant.

Frequently it is desirable to have constitutive or inducible expression of a DNA sequence in particular tissues or organs of a plant. For example, increased nutritional value of a plant might be accomplished by genetic manipulation of the  
25 plant's genome to comprise a seed-preferred promoter operably linked to a heterologous gene such that proteins with enhanced amino acid content are produced in the seed of the plant.

Alternatively, it might be desirable to inhibit expression of a native DNA sequence within a plant's tissues to achieve a desired phenotype. In this case,  
30 such inhibition might be accomplished with transformation of the plant to comprise a tissue-specific promoter operably linked to an antisense nucleotide sequence, such that expression of the antisense sequence produces an RNA transcript that interferes with translation of the mRNA of the native DNA sequence.

Seed development involves embryogenesis and maturation events as well as physiological adaptation processes that occur within the seed to insure progeny survival. Developing plant seeds accumulate and store carbohydrate, lipid, and protein that are subsequently used during germination. Expression of storage 5 protein genes in seeds occurs primarily in the embryonic axis and cotyledons and in the endosperm of developing seeds but never in mature vegetative tissues. Generally, the expression patterns of seed proteins are highly regulated. This regulation includes spatial and temporal regulation during seed development. A variety of proteins accumulate and decay during embryogenesis and seed 10 development and provide an excellent system for investigating different aspects of gene regulation as well as for providing regulatory sequences for use in genetic manipulation of plants.

Thus, isolation and characterization of seed-preferred promoters that can serve as regulatory regions for expression of heterologous nucleotide sequences 15 of interest in a seed-preferred manner are needed for genetic manipulation of plants.

#### SUMMARY OF THE INVENTION

It is an object of the present invention to provide a novel nucleotide 20 sequence for modulating gene expression in a plant.

It is a further object of the present invention to provide an isolated promoter capable of driving transcription in a seed-preferred manner.

It is a further object of the present invention to provide a method of improved control of an endogenous or exogenous product in the seed of a 25 transformed plant.

It is a further object of the present invention to provide a method for providing useful changes in the phenotype of a seed of a transformed plant.

It is a further object of the present invention to provide a method for producing a novel product in the seed of a transformed plant.

30 It is a further object of the present invention to provide a method for producing a novel function in the seed of a transformed plant.

Therefore, in one aspect, the present invention relates to an isolated nucleic acid comprising a member selected from the group consisting of:

- a) sequences capable of driving expression of coding regions selected from the group consisting of coding regions for Cim 1 (cytokinin-induced message); cZ19B1 (maize 19KDa zein); or mi1ps (myo-inositol-1-phosphate synthase);
- 5 b) a sequence comprising at least 20 contiguous nucleotides of the sequence set forth in either of SEQ ID NOS 1, 4, or 7;
- c) a sequence comprising a variant or fragment of the nucleotide sequence set forth in either of SEQ ID NOS 1, 4, or 7;
- d) the nucleotide sequences set forth in SEQ ID NOS 1, 4, or 7;
- 10 e) nucleic acids having at least 60% sequence identity to SEQ ID NO 1, at least 81% sequence identity to SEQ ID NO 4, or at least 60% sequence identity to SEQ ID NO 7, wherein the % sequence identity is based on the entire sequence and is determined by GAP version 10 analysis using default parameters;
- 15 f) a sequence that hybridizes to any one of SEQ ID NOS: 1, 4, or 7, under stringent conditions; and
- g) a nucleic acid complementary to a nucleic acid of (a) through (f).
- 20 In other aspects, the present invention relates to expression cassettes comprising the promoter operably linked to a nucleotide sequence, vectors containing the expression cassette, and plants stably transformed with at least one expression cassette.
- In a further aspect, the present invention relates to a method for modulating expression in the seed of a stably transformed plant comprising the steps of (a) transforming a plant cell with an expression cassette comprising the promoter of the present invention operably linked to at least one nucleotide sequence; (b) growing the plant cell under plant growing conditions and (c) regenerating a stably transformed plant from the plant cell wherein expression of the nucleotide sequence alters the phenotype of the seed.
- 25 Compositions and methods for regulating expression of heterologous nucleotide sequences in a plant are provided. Compositions are novel nucleotide sequences for seed-preferred plant promoters, more particularly transcriptional initiation regions isolated from the plant genes Cim1 (cytokinin-induced message);
- 30 cZ19B1 (maize 19KDa zein); and mi1ps (myo-inositol-1-phosphate synthase). A
- 35

method for expressing a heterologous nucleotide sequence in a plant using the transcriptional initiation sequences disclosed herein is provided. The method comprises transforming a plant cell with a transformation vector that comprises a heterologous nucleotide sequence operably linked to one of the plant promoters of 5 the present invention and regenerating a stably transformed plant from the transformed plant cell. In this manner, the promoter sequences are useful for controlling the expression of endogenous as well as exogenous products in a seed-preferred manner.

Downstream from and under the transcriptional initiation regulation of the 10 seed-specific region will be a sequence of interest which will provide for modification of the phenotype of the seed. Such modification includes modulating the production of an endogenous product, as to amount, relative distribution, or the like, or production of an exogenous expression product to provide for a novel function or product in the seed.

15

#### DETAILED DESCRIPTION OF THE INVENTION

In accordance with the invention nucleotide constructs are provided that allow initiation of transcription in seed. Constructs of the invention comprise regulated transcriptional initiation regions associated with seed formation and 20 seed tissues. Thus, the compositions of the present invention comprise novel nucleotide sequences for plant promoters, more particularly seed-preferred promoters for the genes Cim 1 (cytokinin-induced message); cZ19B1 (maize 19KDa zein); and mi1ps (myo-inositol-1-phosphate synthase).

The promoters for these genes may be isolated from the 5' untranslated 25 region flanking their respective transcription initiation sites. Methods for isolation of promoter regions are well known in the art. By "isolated" is intended that the promoter sequences can be extracted by molecular techniques or synthesized by chemical means. In either instance, the promoter is removed from at least one of its flanking sequences in its native state. Sequences for the promoter regions are 30 set forth as noted above.

Methods are readily available in the art for the hybridization of nucleic acid sequences. Promoter sequences from other plants may be isolated according to well known techniques based on their sequence homology to the promoter sequences set forth herein. In these techniques, all or part of the known promoter

sequence is used as a probe which selectively hybridizes to other sequences present in a population of cloned genomic DNA fragments (i.e. genomic libraries) from a chosen organism.

For example, the entire promoter sequence or portions thereof may be  
5 used as probes capable of specifically hybridizing to corresponding promoter sequences. To achieve specific hybridization under a variety of conditions, such probes include sequences that are unique and are preferably at least about 10 nucleotides in length, and most preferably at least about 20 nucleotides in length. Such probes may be used to amplify corresponding promoter sequences from a  
10 chosen organism by the well-known process of polymerase chain reaction (PCR). This technique may be used to isolate additional promoter sequences from a desired organism or as a diagnostic assay to determine the presence of the promoter sequence in an organism.

Such techniques include hybridization screening of plated DNA libraries  
15 (either plaques or colonies; see e.g. Innis et al. (1990) *PCR Protocols, A Guide to Methods and Applications*, eds., Academic Press).

The terms "stringent conditions" or "stringent hybridization conditions" includes reference to conditions under which a probe will hybridize to its target sequence, to a detectably greater degree than other sequences (e.g., at least 2-fold over background). Stringent conditions are sequence-dependent and will be different in different circumstances. By controlling the stringency of the hybridization and/or washing conditions, target sequences can be identified which are 100% complementary to the probe (homologous probing). Alternatively, stringency conditions can be adjusted to allow some mismatching in sequences so  
20 that lower degrees of similarity are detected (heterologous probing). Generally, a probe is less than about 1000 nucleotides in length, preferably less than 500 nucleotides in length.  
25

Typically, stringent conditions will be those in which the salt concentration is less than about 1.5 M Na ion, typically about 0.01 to 1.0 M Na ion concentration  
30 (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes (e.g., 10 to 50 nucleotides) and at least about 60°C for long probes (e.g., greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamid . Exemplary low stringency conditions include hybridization with a buffer solution of 30 to 35%

formamide, 1 M NaCl, 1% SDS (sodium dodecyl sulphate) at 37°C, and a wash in 1X to 2X SSC (20X SSC = 3.0 M NaCl/0.3 M trisodium citrate) at 50 to 55°C. Exemplary moderate stringency conditions include hybridization in 40 to 45% formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 0.5X to 1X SSC at 55 to 5 60°C. Exemplary high stringency conditions include hybridization in 50% formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 0.1X SSC at 60 to 65°C. Hybridization times may range from about four hours to about sixteen hours and do not define the stringency of the protocol.

Specificity is typically the function of post-hybridization washes, the critical 10 factors being the ionic strength and temperature of the final wash solution. For DNA-DNA hybrids, the  $T_m$  can be approximated from the equation of Meinkoth and Wahl, Anal. Biochem., 138:267-284 (1984):  $T_m = 81.5^{\circ}\text{C} + 16.6(\log M) + 0.41(\%GC) - 0.61(\% \text{form}) - 500/L$ ; where M is the molarity of monovalent cations, %GC is the percentage of guanosine and cytosine nucleotides in the DNA, % form 15 is the percentage of formamide in the hybridization solution, and L is the length of the hybrid in base pairs. The  $T_m$  is the temperature (under defined ionic strength and pH) at which 50% of a complementary target sequence hybridizes to a perfectly matched probe.  $T_m$  is reduced by about 1°C for each 1% of mismatching; thus,  $T_m$ , hybridization and/or wash conditions can be adjusted to 20 hybridize to sequences of the desired identity. For example, if sequences with >90% identity are sought, the  $T_m$  can be decreased 10°C. Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point ( $T_m$ ) for the specific sequence and its complement at a defined ionic strength and pH. However, severely stringent conditions can utilize a hybridization and/or wash at 1, 25 2, 3, or 4°C lower than the thermal melting point ( $T_m$ ); moderately stringent conditions can utilize a hybridization and/or wash at 6, 7, 8, 9, or 10°C lower than the thermal melting point ( $T_m$ ); low stringency conditions can utilize a hybridization and/or wash at 11, 12, 13, 14, 15, or 20°C lower than the thermal melting point ( $T_m$ ). Using the equation, hybridization and wash compositions, and desired  $T_m$ , 30 those of ordinary skill will understand that variations in the stringency of hybridization and/or wash solutions are inherently described. For purposes of defining the invention, stringent conditions will be that which includes a wash at about 5°C lower than the thermal melting point ( $T_m$ ) for the specific sequence and its complement at a defined ionic strength and pH. If the desired degree of

- mismatching results in a  $T_m$  of less than 45°C (aqueous solution) or 32°C (formamide solution) it is preferred to increase the SSC concentration so that a higher temperature can be used. An extensive guide to the hybridization of nucleic acids is found in Tijssen, *Laboratory Techniques in Biochemistry and Molecular*
- 5     Biology--Hybridization with Nucleic Acid Probes, Part I, Chapter 2 "Overview of principles of hybridization and the strategy of nucleic acid probe assays", Elsevier, New York (1993); and Current Protocols in Molecular Biology, Chapter 2, Ausubel, et al., Eds., Greene Publishing and Wiley-Interscience, New York (1995). See Sambrook et al. (1989) *Molecular Cloning: A Laboratory Manual* (2nd ed. Cold
- 10    Spring Harbor Laboratory, Cold Spring Harbor, N.Y.) In general, sequences that correspond to promoter sequences of the invention and hybridize to the promoter sequence disclosed herein will be at least 50% homologous, 70% homologous, and even 85% homologous or more with the disclosed sequences. That is, the sequence similarity of sequences may range, sharing at least about 50%, about
- 15    70%, and even about 85% sequence similarity.

The promoter regions of the invention may be isolated from any plant, including, but not limited to corn (*Zea mays*), non-vegetable *Brassica* (*Brassica napus*, *Brassica rapa* ssp.), alfalfa (*Medicago sativa*), rice (*Oryza sativa*), rye (*Secale cereale*), sorghum (*Sorghum bicolor*, *Sorghum vulgare*), sunflower (*Helianthus annuus*), wheat (*Triticum aestivum*), soybean (*Glycine max*), tobacco (*Nicotiana tabacum*), potato (*Solanum tuberosum*), peanuts (*Arachis hypogaea*), cotton (*Gossypium hirsutum*), sweet potato (*Ipomoea batatas*), cassava (*Manihot esculenta*), coffee (*Coffea spp.*), coconut (*Cocos nucifera*), pineapple (*Ananas comosus*), citrus trees (*Citrus spp.*), cocoa (*Theobroma cacao*), tea (*Camellia sinensis*), banana (*Musa spp.*), avocado (*Persea americana*), fig (*Ficus casica*), guava (*Psidium guajava*), mango (*Mangifera indica*), olive (*Olea europaea*), oats, barley, vegetables, ornamentals, and conifers. Preferably, plants include corn, soybean, sunflower, safflower, *Brassica*, wheat, barley, rye, alfalfa, and sorghum.

The coding sequence expressed by the promoters of the invention may be used for varying the phenotype of the seeds. Various changes in phenotype are of interest including modifying the fatty acid composition in seeds, altering the starch or carbohydrate profile, altering the amino acid content of the seed, and the like. These results can be achieved by providing expression of heterologous or increased expression of endogenous products in seeds. Alternatively, the results

can be achieved by providing for a reduction of expression of one or more endogenous products, particularly enzymes or cofactors in the seed. These changes result in a change in ph notype of the transformed seed.

- Genes of interest include, generally, those involved in oil, starch,  
5 carbohydrate or nutrient metabolism as well as those affecting kernel size,  
sucrose loading, and the like.

General categories of genes of interest for the purposes of the present invention include for example, those genes involved in information, such as Zinc fingers, those involved in communication, such as kinases, and those involved in  
10 housekeeping, such as heat shock proteins. More specific categories of transgenes, for example, include genes encoding important traits for agronomics, insect resistance, disease resistance, herbicide resistance, and grain characteristics. It is recognized that any gene of interest can be operably linked to the promoter of the invention and expressed in the seed.

15 Agronomically important traits such as oil, starch and protein content can be genetically altered in addition to using traditional breeding methods. Modifications include increasing content of oleic acid, saturated and unsaturated oils, increasing levels of lysine and sulfur and providing essential amino acids, and also modification of starch. Hordeothionin protein modifications are described in  
20 U.S. Serial Nos. 08/838,763 filed April 10, 1997; 08/824,379 filed March 26, 1997; 08/824,382 filed March 26, 1997; and U.S. Pat. No. 5,703,409 issued December 30, 1997, the disclosures of which are incorporated herein in their entirety by reference. Another example is lysine and/or sulfur rich seed protein encoded by the soybean 2S albumin described in U.S. Serial No. 08/618,911 filed March 20,  
25 1996, and the chymotrypsin inhibitor from barley, Williamson *et al.* (1987) *Eur. J. Biochem.* 165:99-106, the disclosures of each are incorporated by reference. Derivatives of the following genes can be made by site directed mutagenesis to increase the level of preselected amino acids in the encoded polypeptide. For example, the gene encoding the barley high lysine polypeptide (BHL), is derived  
30 from barley chymotrypsin inhibitor, U.S. Serial No. 08/740,682 filed November 1, 1996 and PCT/US97/20441 filed October 31, 1997, the disclosures of each are incorporated herein by reference. Other proteins include methionine-rich plant proteins such as from sunflower seed (Lilley *et al.* (1989) *Proceedings of the World Congress on Vegetable Protein Utilization in Human Foods and Animal*

- Feedstuffs; Applewhite, H. (ed.); American Oil Chemists Soc., Champaign, IL:497-502, incorporated herein in its entirety by reference), corn (Pedersen *et al.* (1986) *J. Biol. Chem.* 261:6279; Kiriha *et al.* (1988) *Gene* 71:359, both incorporated herein in its entirety by reference) and rice (Musumura *et al.* (1989) *Plant Mol. Biol.* 12:123, incorporated herein in its entirety by reference). Other agronomically important genes encode latex, Flory 2, growth factors, seed storage factors and transcription factors.
- Agronomic traits in seeds can be improved by altering expression of genes that affect the response of seed growth and development during environmental stress, Cheikh-N *et al.* (1994) *Plant Physiol.* 106(1):45-51) and genes controlling carbohydrate metabolism to reduce kernel abortion in maize, Zinselmeier *et al.* (1995) *Plant Physiol.* 107(2):385-391.
- Insect resistance genes may encode resistance to pests that have great yield drag such as rootworm, cutworm, European Corn Borer, and the like. Such genes include, for example, *Bacillus thuringiensis* endotoxin genes (U.S. Pat. Nos. 5,366,892; 5,747,450; 5,737,514; 5,723,756; 5,593,881; Geiser *et al.* (1986) *Gene* 48:109); lectins (Van Damme *et al.* (1994) *Plant Mol. Biol.* 24:825); and the like.
- Genes encoding disease resistance traits may include detoxification genes, such as against fumonosin (U.S. Patent Application No. 08/484,815 filed June 7, 1995); avirulence (avr) and disease resistance (R) genes (Jones *et al.* (1994) *Science* 266:789; Martin *et al.* (1993) *Science* 262:1432; Mindrinos *et al.* (1994) *Cell* 78:1089; and the like.
- The quality of grain is reflected in traits such as levels and types of oils, saturated and unsaturated, quality and quantity of essential amino acids, and levels of cellulose. In corn, modified hordothionin proteins, described in U.S. Serial Nos. 08/838,763 filed April 10, 1997; 08/824,379 filed March 26, 1997; 08/824,382 filed March 26, 1997; and U.S. Pat. No. 5,703,409 issued December 30, 1997, provide descriptions of modifications of proteins for desired purposes.
- Commercial traits can also be encoded on a gene(s) which could alter or increase for example, starch for the production of paper, textiles and ethanol, or provide expression of proteins with other commercial uses. Another important commercial use of transformed plants is the production of polymers and bioplastics such as described in U.S. Patent No. 5,602,321 issued February 11, 1997. Genes such as  $\beta$ -K tothiolase, PHBase (polyhydroxybutyrate synthase)

and acetoacetyl-CoA reductase (see Schubert *et al.* (1988) *J. Bacteriol.* 170(12):5837-5847) facilitate expression of polyhydroxyalkanoates (PHAs).

Exogenous products include plant enzymes and products as well as those from other sources including prokaryotes and other eukaryotes. Such products 5 include enzymes, cofactors, hormones, and the like. The level of seed proteins, particularly modified seed proteins having improved amino acid distribution to improve the nutrient value of the seed can be increased. This is achieved by the expression of such proteins having enhanced amino acid content.

As noted, the heterologous nucleotide sequence operably linked to one of 10 the promoters disclosed herein may be an antisense sequence for a targeted gene. By "antisense DNA nucleotide sequence" is intended a sequence that is in inverse orientation to the 5'-to-3' normal orientation of that nucleotide sequence. When delivered into a plant cell, expression of the antisense DNA sequence prevents normal expression of the DNA nucleotide sequence for the targeted 15 gene. The antisense nucleotide sequence encodes an RNA transcript that is complementary to and capable of hybridizing to the endogenous messenger RNA (mRNA) produced by transcription of the DNA nucleotide sequence for the targeted gene. In this case, production of the native protein encoded by the targeted gene is inhibited to achieve a desired phenotypic response. Thus the 20 promoter sequences disclosed herein may be operably linked to antisense DNA sequences to reduce or inhibit expression of a native protein in the plant seed.

By "promoter" or "transcriptional initiation region" is intended a regulatory region of DNA usually comprising a TATA box capable of directing RNA polymerase II to initiate RNA synthesis at the appropriate transcription initiation 25 site for a particular coding sequence. A promoter may additionally comprise other recognition sequences generally positioned upstream or 5' to the TATA box, referred to as upstream promoter elements, which influence the transcription initiation rate. It is recognized that having identified the nucleotide sequences for the promoter regions disclosed herein, it is within the state of the art to isolate and 30 identify further regulatory elements in the 5' untranslated region upstream from the particular promoter regions identified herein. Thus the promoter regions disclosed herein are generally further defined by comprising upstream regulatory elements such as those responsible for tissue and temporal expression of the coding sequence, enhancers and the like. In the same manner, the promoter elements

which enable expression in the desired tissue such as the seed can be identified, isolated, and used with other core promoters to confirm seed-preferred expression.

The regulatory sequences of the present invention, when operably linked to  
5 a heterologous nucleotide sequence of interest and inserted into a transformation vector, enable seed-preferred expression of the heterologous nucleotide sequence in the seeds of a plant stably transformed with this vector. By "seed-preferred" is intended expression in the seed, including at least one of embryo, kernel, pericarp, endosperm, nucellus, aleurone, pedicel, and the like.

10 By "heterologous nucleotide sequence" is intended a sequence that is not naturally occurring with the promoter sequence. While this nucleotide sequence is heterologous to the promoter sequence, it may be homologous, or native, or heterologous, or foreign, to the plant host.

15 It is recognized that the promoters may be used with their native coding sequences to increase or decrease expression resulting in a change in phenotype in the transformed seed.

The isolated promoter sequences of the present invention can be modified to provide for a range of expression levels of the heterologous nucleotide sequence. Thus, less than the entire promoter regions may be utilized and the  
20 ability to drive seed-preferred expression retained. However, it is recognized that expression levels of mRNA may be decreased with deletions of portions of the promoter sequences. Generally, at least about 20 nucleotides of an isolated promoter sequence will be used to drive expression of a nucleotide sequence.

It is recognized that to increase transcription levels enhancers may be  
25 utilized in combination with the promoter regions of the invention. Enhancers are nucleotide sequences that act to increase the expression of a promoter region. Enhancers are known in the art and include the SV40 enhancer region, the 35S enhancer element, and the like.

Modifications of the isolated promoter sequences of the present invention  
30 can provide for a range of expression of the heterologous nucleotide sequence. Thus, they may be modified to be weak promoters or strong promoters. Generally, by "weak promoter" is intended a promoter that drives expression of a coding sequence at a low level. By "low level" is intended at levels of about 1/10,000 transcripts to about 1/100,000 transcripts to about 1/500,000 transcripts.

conversely, a strong promoter drives expression of a coding sequence at a high level, or at about 1/10 transcripts to about 1/100 transcripts to about 1/1,000 transcripts.

- The nucleotide sequences for the promoters of the present invention may
- 5 be the naturally occurring sequences or sequences having substantial homology. By "substantial homology" is intended a sequence exhibiting substantial functional and structural equivalence with the naturally occurring sequence. Any structural differences between substantially homologous sequences do not effect the ability of the sequence to function as a promoter as disclosed in the present invention.
- 10 Thus, sequences having substantial sequence homology with the sequence of a particular seed-preferred promoter of the present invention will direct seed-preferred expression of an operably linked heterologous nucleotide sequence. Two promoter nucleotide sequences are considered substantially homologous when they have at least about 70%, preferably at least about 80%, more
- 15 preferably at least about 90%, still more preferably at least about 95% sequence homology. Substantially homologous sequences of the present invention include variants of the disclosed sequences such as those that result from site-directed mutagenesis, as well as synthetically derived sequences.

Substantially homologous sequences of the present invention also refer to

20 those fragments of a particular promoter nucleotide sequence disclosed herein that operate to promote the seed-preferred expression of an operably linked heterologous nucleotide sequence. These fragments will comprise at least about 20 contiguous nucleotides, preferably at least about 50 contiguous nucleotides, more preferably at least about 75 contiguous nucleotides, even more preferably at

25 least about 100 contiguous nucleotides of the particular promoter nucleotide sequence disclosed herein. The nucleotides of such fragments will usually comprise the TATA recognition sequence of the particular promoter sequence. Such fragments may be obtained by use of restriction enzymes to cleave the naturally occurring promoter nucleotide sequences disclosed herein; by

30 synthesizing a nucleotide sequence from the naturally occurring promoter DNA sequence; or may be obtained through the use of PCR technology. See particularly, Mullis *et al.* (1987) *Methods Enzymol.* 155:335-350, and Erlich, ed. (1989) *PCR Technology* (Stockton Press, New York). Again, variants of these

promoter fragments, such as those resulting from site-directed mutagenesis, are encompassed by the compositions of the present invention.

Nucleotide sequences comprising at least about 20 contiguous sequences of the sequences set forth in SEQ ID NOS:1, 4, or 7 are encompassed. These 5 sequences may be isolated by hybridization, PCR, and the like. Such sequences encompass fragments capable of driving seed-preferred expression, fragments useful as probes to identify similar sequences, as well as elements responsible for temporal or tissue specificity. Biologically active variants of the promoter sequences are also encompassed by the method of the present invention. Such 10 variants should retain promoter activity, particularly the ability to drive expression in seed or seed tissues. Biologically active variants include, for example, the native promoter sequences of the invention having one or more nucleotide substitutions, deletions or insertions. Promoter activity may be measured by Northern blot analysis, reporter activity measurements when using transcriptional 15 fusions, and the like. See, for example, Sambrook *et al.* (1989) *Molecular Cloning: A Laboratory Manual* (2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.), herein incorporated by reference.

The following terms are used to describe the sequence relationships between two or more nucleic acids or polynucleotides: (a) "reference sequence", 20 (b) "comparison window", (c) "percentage of sequence identity", and (d) "substantial identity".

(a) As used herein, "reference sequence" is a defined sequence used as a basis for sequence comparison. A reference sequence may be a subset or the entirety of a specified sequence; for example, as a segment of a full-length promoter sequence, or the complete promoter sequence.

(b) As used herein, "comparison window" makes reference to a contiguous and specified segment of a polynucleotide sequence, wherein the polynucleotide sequence may be compared to a reference sequence and wherein the portion of the polynucleotide sequence in the comparison window may 30 comprise additions or deletions (i.e., gaps) compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. Generally, the comparison window is at least 20 contiguous nucleotides in length and optionally can be 30, 40, 50, 100, or more contiguous nucleotides in length. Those of skill in the art understand that to avoid a high

similarity to a reference sequence due to inclusion of gaps in the polynucleotide sequence a gap penalty is typically introduced and is subtracted from the number of matches.

- Methods of alignment of sequences for comparison are well-known in the art. Optimal alignment of sequences for comparison may be conducted by the homology alignment algorithm of Needleman and Wunsch (1970) J. Mol. Biol. 48:443. Computerized implementation of this algorithm includes, but is not limited to GAP, and BLAST, in the Wisconsin Genetics Software Package, Genetics Computer Group (GCG) (575 Science Drive, Madison, Wisconsin. An example of the BLAST family of programs, which can be used to search database sequence similarity for the purposes of this invention, includes BLASTN program for nucleotide query sequences against nucleotide sequence dataset. See, Ausubel et al., eds. (1995) Current Protocols in Molecular Biology, Chapter 19 (Greene Publishing and Wiley-Interscience, New York). The BLAST homology alignment algorithm is useful for comparing fragments of the reference nucleotide or amino acid sequence to sequences from public databases. It is then necessary to apply a method of aligning the complete reference sequence to the database sequences to establish a percentage of identity (in the case of polynucleotides) or a percentage of similarity (in the case of polypeptides). The GAP algorithm is such a method.

GAP uses the algorithm of Needleman and Wunsch (J. Mol. Biol. 48: 443-453, 1970) to find the alignment of two complete sequences that maximizes the number of matches and minimizes the number of gaps. GAP considers all possible alignments and gap positions and creates the alignment with the largest number of matched bases and the fewest gaps. It allows for the provision of a gap creation penalty and a gap extension penalty in units of matched bases. GAP must make a profit of gap creation penalty number of matches for each gap it inserts. If a gap extension penalty greater than zero is chosen, GAP must, in addition, make a profit for each gap inserted of the length of the gap times the gap extension penalty. Default gap creation penalty values and gap extension penalty values in Version 10 of the Wisconsin Genetics Software Package for protein sequences are 8 and 2, respectively. For nucleotide sequences the default gap creation penalty is 50 while the default gap extension penalty is 3. The gap creation and gap extension penalties can be expressed as an integer selected

from the group of integers consisting of from 0 to 200. Thus, for example, the gap creation and gap extension penalties can be 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65 or greater.

GAP presents one member of the family of best alignments. There may be  
5 many members of this family, but no other member has a better quality. GAP displays four figures of merit for alignments: Quality, Ratio, Identity, and Similarity. The Quality is the metric maximized in order to align the sequences. Ratio is the quality divided by the number of bases in the shorter segment. Percent Identity is the percent of the symbols that actually match. Percent Similarity is the percent of  
10 the symbols that are similar. Symbols that are across from gaps are ignored. A similarity is scored when the scoring matrix value for a pair of symbols is greater than or equal to 0.50, the similarity threshold. The scoring matrix used in Version 10 of the Wisconsin Genetics Software Package is BLOSUM62 (see Henikoff & Henikoff (1989) Proc. Natl. Acad. Sci. USA 89:10915).

15 Unless otherwise stated, for purposes of the invention, the preferred method of determining percent sequence identity is by the GAP version 10 algorithm using default parameters.

(c) As used herein, "percentage of sequence identity" means the value determined by comparing two optimally aligned sequences over a  
20 comparison window, wherein the portion of the polynucleotide sequence in the comparison window may comprise additions or deletions (i.e., gaps) as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid base  
25 occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison and multiplying the result by 100 to yield the percentage of sequence identity.

(d) The term "substantial identity" of polynucleotide sequences  
30 means that a polynucleotide comprises a sequence that has at least 70% sequence identity, preferably at least 80%, more preferably at least 90% and most preferably at least 95%, compared to a reference sequence using one of the alignment programs described using standard parameters.

Another indication that nucleotide sequences are substantially identical is if two nucleic acid molecules hybridize to each other under stringent conditions. Generally, stringent temperature conditions are selected to be about 5°C to about 2°C lower than the melting point ( $T_m$ ) for the specific sequence at a defined 5 ionic strength and pH. The denaturation or melting of DNA occurs over a narrow temperature range and represents the disruption of the double helix into its complementary single strands. The process usually is characterized by the temperature of the midpoint of transition,  $T_m$ , which is sometimes described as the melting temperature. Formulas are available in the art for the determination of 10 melting temperatures. Typically, stringent wash conditions are those in which the salt concentration is about 0.02 molar at pH 7 and the temperature is at 50, 55, or 60°C.

The nucleotide sequences for the seed-preferred promoters disclosed in the present invention, as well as variants and fragments thereof, are useful in the 15 genetic manipulation of any plant when operably linked with a heterologous nucleotide sequence whose expression is to be controlled to achieve a desired phenotypic response. By "operably linked" is intended the transcription or translation of the heterologous nucleotide sequence is under the influence of the promoter sequence. In this manner, the nucleotide sequences for the promoters of 20 the invention may be provided in expression cassettes along with heterologous nucleotide sequences for expression in the plant of interest, more particularly in the seed of the plant.

Such expression cassettes will comprise a transcriptional initiation region comprising one of the promoter nucleotide sequences of the present invention, or 25 variants or fragments thereof, operably linked to the heterologous nucleotide sequence whose expression is to be controlled by the seed-preferred promoters disclosed herein. Such an expression cassette is provided with a plurality of restriction sites for insertion of the nucleotide sequence to be under the transcriptional regulation of the regulatory regions. The expression cassette may 30 additionally contain selectable marker genes.

The expression cassette will include in the 5'-to-3' direction of transcription, a transcriptional and translational initiation region, a heterologous nucleotide sequence of interest, and a transcriptional and translational termination region functional in plants. The termination region may be native with the transcriptional

initiation region comprising one of the promoter nucleotide sequences of the present invention, may be native with the DNA sequence of interest, or may be derived from another source. Convenient termination regions are available from the Ti-plasmid of *A. tumefaciens*, such as the octopine synthase and nopaline

5 synthase termination regions. See also, Guerineau *et al.* (1991) *Mol. Gen. Genet.* 262:141-144; Proudfoot (1991) *Cell* 64:671-674; Sanfacon *et al.* (1991) *Genes Dev.* 5:141-149; Mogen *et al.* (1990) *Plant Cell* 2:1261-1272; Munroe *et al.* (1990) *Gene* 91:151-158; Ballas *et al.* 1989) *Nucleic Acids Res.* 17:7891-7903; Joshi *et al.* (1987) *Nucleic Acid Res.* 15:9627-9639.

10 The expression cassette comprising the promoter sequence of the present invention operably linked to a heterologous nucleotide sequence may also contain at least one additional nucleotide sequence for a gene to be cotransformed into the organism. Alternatively, the additional sequence(s) can be provided on another expression cassette.

15 Where appropriate, the heterologous nucleotide sequence whose expression is to be under the control of the promoter sequence of the present invention and any additional nucleotide sequence(s) may be optimized for increased expression in the transformed plant. That is, these nucleotide sequences can be synthesized using plant preferred codons for improved 20 expression. Methods are available in the art for synthesizing plant-preferred nucleotide sequences. See, for example, U.S. Patent Nos. 5,380,831 and 5,436,391, and Murray *et al.* (1989) *Nucleic Acids Res.* 17:477-498, herein incorporated by reference.

Additional sequence modifications are known to enhance gene expression 25 in a cellular host. These include elimination of sequences encoding spurious polyadenylation signals, exon-intron splice site signals, transposon-like repeats, and other such well-characterized sequences that may be deleterious to gene expression. The G-C content of the heterologous nucleotide sequence may be adjusted to levels average for a given cellular host, as calculated by reference to 30 known genes expressed in the host cell. When possible, the sequence is modified to avoid predicted hairpin secondary mRNA structures.

The expression cassettes may additionally contain 5' leader sequences in the expression cassette construct. Such leader sequences can act to enhance translation. Translation leaders are known in the art and include: picornavirus

leaders, for example, EMCV leader (Encephalomyocarditis 5' noncoding region) (Elroy-Stein *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86:6126-6130); potyvirus leaders, for example, TEV leader (Tobacco Etch Virus) (Allison *et al.* (1986)); MDMV leader (Maize Dwarf Mosaic Virus) (*Virology* 154:9-20); human immunoglobulin heavy-chain binding protein (BiP) (Macejak *et al.* (1991) *Nature* 353:90-94); untranslated leader from the coat protein mRNA of alfalfa mosaic virus (AMV RNA 4) (Jobling *et al.* (1987) *Nature* 325:622-625); tobacco mosaic virus leader (TMV) (Gallie *et al.* (1989) *Molecular Biology of RNA*, pages 237-256); and maize chlorotic mottle virus leader (MCMV) (Lommel *et al.* (1991) *Virology* 81:382-385). See also Della-Cioppa *et al.* (1987) *Plant Physiology* 84:965-968. Other methods known to enhance translation and/or mRNA stability can also be utilized, for example, introns, and the like.

In those instances where it is desirable to have the expressed product of the heterologous nucleotide sequence directed to a particular organelle, particularly the plastid, amyloplast or vacuole, or to the endoplasmic reticulum, or secreted at the cell's surface or extracellularly, the expression cassette may further comprise a coding sequence for a transit peptide. Such transit peptides are well known in the art and include, but are not limited to, the transit peptide for the acyl carrier protein, the small subunit of RUBISCO, plant EPSP synthase, and the like.

In preparing the expression cassette, the various DNA fragments may be manipulated, so as to provide for the DNA sequences in the proper orientation and, as appropriate, in the proper reading frame. Toward this end, adapters or linkers may be employed to join the DNA fragments or other manipulations may be involved to provide for convenient restriction sites, removal of superfluous DNA, removal of restriction sites, or the like. For this purpose, in vitro mutagenesis, primer repair, restriction, annealing, resubstitutions, for example, transitions and transversions, may be involved.

Reporter genes or selectable marker genes may be included in the expression cassettes. Examples of suitable reporter genes known in the art can be found in, for example, Jefferson *et al.* (1991) in *Plant Molecular Biology Manual*, ed. Gelvin *et al.* (Kluwer Academic Publishers), pp. 1-33; DeWet *et al.* (1987) *Mol. Cell. Biol.* 7:725-737; Goff *et al.* (1990) *EMBO J.* 9:2517-2522; Kain *et*

al. (1995) *BioTechniques* 19:650-655; and Chiu et al. (1996) *Current Biology* 6:325-330.

Selectable marker genes for selection of transformed cells or tissues can include genes that confer antibiotic resistance or resistance to herbicides.

- 5 Examples of suitable selectable marker genes include, but are not limited to, genes encoding resistance to chloramphenicol (Herrera Estrella et al. (1983) *EMBO J.* 2:987-992); methotrexate (Herrera Estrella et al. (1983) *Nature* 303:209-213; Meijer et al. (1991) *Plant Mol. Biol.* 16:807-820); hygromycin (Waldron et al. (1985) *Plant Mol. Biol.* 5:103-108; Zhijian et al. (1995) *Plant Science* 108:219-10 227); streptomycin (Jones et al. (1987) *Mol. Gen. Genet.* 210:86-91); spectinomycin (Bretagne-Sagnard et al. (1996) *Transgenic Res.* 5:131-137); bleomycin (Hille et al. (1990) *Plant Mol. Biol.* 7:171-176); sulfonamide (Guerineau et al. (1990) *Plant Mol. Biol.* 15:127-136); bromoxynil (Stalker et al. (1988) *Science* 242:419-423); glyphosate (Shaw et al. (1986) *Science* 233:478-481); 15 phosphinothricin (DeBlock et al. (1987) *EMBO J.* 6:2513-2518).

Other genes that could serve utility in the recovery of transgenic events but might not be required in the final product would include, but are not limited to, examples such as GUS (glucuronidase; Jefferson (1987) *Plant Mol. Biol. Rep.* 5:387), GFP (green fluorescence protein; Chalfie et al. (1994) *Science* 263:802), 20 luciferase (Riggs et al. (1987) *Nucleic Acids Res.* 15(19):8115; Luehrs et al. (1992) *Methods Enzymol.* 216:397-414) and the maize genes encoding for anthocyanin production (Ludwig et al. (1990) *Science* 247:449).

Transformation protocols as well as protocols for introducing nucleotide sequences into plants may vary depending on the type of plant or plant cell, i.e., 25 monocot or dicot, targeted for transformation. Suitable methods of introducing nucleotide sequences into plant cells and subsequent insertion into the plant genome include microinjection (Crossway et al. (1986) *Biotechniques* 4:320-334), electroporation (Riggs et al. (1986) *Proc. Natl. Acad. Sci. USA* 83:5602-5606, Agrobacterium-mediated transformation (Townsend et al., U.S. Pat No. 5,563,055; 30 Zhao et al WO US98/01268), direct gene transfer (Paszkowski et al. (1984) *EMBO J.* 3:2717-2722), and ballistic particle acceleration (see, for example, Sanford et al., U.S. Patent No. 4,945,050; Tomes et al. (1995) "Direct DNA Transfer into Intact Plant Cells via Microprojectile Bombardment," in *Plant Cell, Tissue, and Organ Culture: Fundamental Methods*, ed. Gamborg and Phillips (Springer-Verlag,

Berlin); and McCabe *et al.* (1988) *Biotechnology* 6:923-926). Also see Weissinger *et al.* (1988) *Ann. Rev. Genet.* 22:421-477; Sanford *et al.* (1987) *Particulate Science and Technology* 5:27-37 (onion); Christou *et al.* (1988) *Plant Physiol.* 87:671-674 (soybean); McCabe *et al.* (1988) *Bio/Technology* 6:923-926

5 (soybean); Finer and McMullen (1991) *In Vitro Cell Dev. Biol.* 27P:175-182 (soybean); Singh *et al.* (1998) *Theor. Appl. Genet.* 96:319-324 (soybean); Datta *et al.* (1990) *Biotechnology* 8:736-740 (rice); Klein *et al.* (1988) *Proc. Natl. Acad. Sci. USA* 85:4305-4309 (maize); Klein *et al.* (1988) *Biotechnology* 6:559-563 (maize); Tomes, U.S. Patent No. 5,240,855; Busing *et al.*, U.S. Patent Nos. 5,322,783 and 10 5,324,646; Tomes *et al.* (1995) "Direct DNA Transfer into Intact Plant Cells via Microprojectile Bombardment," in *Plant Cell, Tissue, and Organ Culture: Fundamental Methods*, ed. Gamborg (Springer-Verlag, Berlin) (maize); Klein *et al.* (1988) *Plant Physiol.* 91:440-444 (maize); Fromm *et al.* (1990) *Biotechnology* 8:833-839 (maize); Hooykaas-Van Slooteren *et al.* (1984) *Nature* (London) 15 311:763-764; Bowen *et al.*, U.S. Patent No. 5,736,369 (cereals); Bytebier *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:5345-5349 (Liliaceae); De Wet *et al.* (1985) in *The Experimental Manipulation of Ovule Tissues*, ed. Chapman *et al.* (Longman, New York), pp. 197-209 (pollen); Kaepler *et al.* (1990) *Plant Cell Reports* 9:415-418 and Kaepler *et al.* (1992) *Theor. Appl. Genet.* 84:560-566

20 (whisker-mediated transformation); D'Halluin *et al.* (1992) *Plant Cell* 4:1495-1505 (electroporation); Li *et al.* (1993) *Plant Cell Reports* 12:250-255 and Christou and Ford (1995) *Annals of Botany* 75:407-413 (rice); Osjoda *et al.* (1996) *Nature Biotechnology* 14:745-750 (maize via *Agrobacterium tumefaciens*); all of which are herein incorporated by reference.

25 The cells that have been transformed may be grown into plants in accordance with conventional ways. See, for example, McCormick *et al.* (1986) *Plant Cell Reports* 5:81-84. These plants may then be grown, and either pollinated with the same transformed strain or different strains, and the resulting hybrid having constitutive expression of the desired phenotypic characteristic identified. Two or more generations may be grown to ensure that constitutive expression of the desired phenotypic characteristic is stably maintained and inherited and then seeds harvested to ensure constitutive expression of the 30 desired phenotypic characteristic has been achieved.

The expression cassette comprising the particular promoter sequence of the present invention operably linked to a heterologous nucleotide sequence of interest can be used to transform any plant. In this manner, genetically modified plants, plant cells, plant tissue, seed, and the like can be obtained.

- 5       The following examples are offered by way of illustration and not by way of limitation.

## EXPERIMENTAL

Promoter regions for the maize genes Cim 1 (cytokinin-induced message) 10 (Genbank Acc. No. U03860); cZ19B1 (maize 19KDa zein) (Genbank Acc. No. M12143); and mi1ps (myo-inositol-1-phosphate synthase) (Genbank Acc. No. U32511); were isolated from maize plants and cloned. These genes were selected as sources of seed-preferred promoters based on the spatial expression of their gene products. The method for their isolation is described below.

15

### Example 1: Isolation of Promoter Sequences

The procedure for promoter isolation is described in the User Manual for the Genome Walker kit sold by Clontech Laboratories, Inc., Palo Alto, California. Genomic DNA from maize line V3-4 A63 was prepared by grinding 10-day-old 20 seedling leaves in liquid nitrogen, and the DNA prepared as described by Chen and Dellaporta (1994) in *The Maize Handbook*, ed. Freeling and Walbot (Springer-Verlag, Berlin) with a few minor modifications. Precipitated DNA was recovered using an inoculation loop and transferred to a 1.5 ml eppendorf tube containing 500 μl of TE(10 mM Tris pH 8.0, 1 mM EDTA). The DNA was allowed to dissolve 25 at room temperature for 15 minutes, phenol extracted and 2-propanol precipitated in 700 μl. The precipitate was recovered and washed with 70% ethanol. The DNA was then placed in a clean 1.5 ml eppendorf tube to air dry and resuspended in 200 μl of TE. RNase A was added to 10 μg/ml and the mixture was incubated at 37°C for several hours. The DNA was then extracted once with phenol-chloroform, then chloroform, then ethanol precipitated and resuspended in TE. 30 The DNA was then used exactly as described in the Genome Walker User Manual (Clontech PT3042-1 version PR68687). Briefly, the DNA was digested separately with restriction enzymes DraI, EcoRV, PvuII, ScaI, and StuI; all blunt-end cutters. The DNA was extracted with phenol, then chloroform, then ethanol precipitated.

The Genome Walker adapt<sup>rs</sup>s were ligated onto the ends of the restrict<sup>d</sup> d DNA. The resulting DNA is referred to as DL1-DL5, respectively.

For isolation of specific promoter regions, two nonoverlapping gene-specific primers (27-30 bp in length) were designed from the 5' end of the maize genes  
5 identified from sequence databases. The primers were designed to amplify the region upstream of the coding sequence, i.e. the 5' untranslated region and promoter of the chosen gene. The sequence of the primers are given below for each promoter described. The first round of PCR was performed on each DNA sample (DL1-5) with Clontech primer AP1 and the gene-specific primers (gsp)1  
10 with the sequences shown in SEQ ID NOS: 1, 5, or 8.

PCR was performed in a model PTC-100 thermal cycler with HotBonnet from MJ Research (Watertown, MA) using reagents supplied with the Genome Walker kit. The following cycle parameters were used: 7 cycles of 94°C for 2 seconds, then 72°C for 3 minutes, followed by 32 cycles of 94°C for 2 seconds  
15 and 67°C for 3 minutes. Finally, the samples were held at 67°C for 4 minutes and then at 4°C until further analysis.

As described in the User Manual, the DNA from the first round of PCR was then diluted and used as a template in a second round of PCR using the Clontech AP2 primer and gene-specific primers (gsp)2 with the sequences shown in SEQ  
20 ID NOS: 3, 6, or 9.

The cycle parameters for the second round were: 5 cycles of 94°C for 2 seconds, then 72°C for 3 minutes. Finally, the samples were held at 67°C for 4 minutes and then held at 4°C. Approximately 10<sup>l</sup> o f each reaction were run on a 0.8% agarose gel, and bands (usually 500bp or larger) were excised, purified with  
25 the Sephadaglas BandPrep kit (Pharmacia, Piscataway, NJ) and cloned into the TOPOTA vector pCR2.1 (Invitrogen, Carlsbad, CA; [www.invitrogen.com](http://www.invitrogen.com)). Clones were sequenced for verification.

#### Example 2: Expression Data Using Promoter Sequences

30 Three promoter::GUS fusion constructs were prepared by the methods described below. All vectors were constructed using standard molecular biology techniques (Sambrook *et al.*, *Supra*). A reporter gene and a selectable marker gene for gene expression and selection was inserted between the multiple cloning sites of the pBluescript cloning vector (Stratagene Inc., 11011 N. Torrey Pines

Rd., La Jolla, CA; [www.strategene.com](http://www.strategene.com)). The reporter gene was the - glucuronidase (GUS) gene (Jefferson, R.A. et al., 1986, *Proc. Natl. Acad. Sci.* (USA) 83:8447-8451) into whose coding region was inserted the second intron from the potato ST-LS1 gene (Vancanneyt et al., Mol. Gen. Genet. 220:245-250, 5 1990), to produce intron-GUS, in order to prevent expression of the gene in Agrobacterium (see Ohta, S. et al., 1990, Plant Cell Physiol. 31(6):805-813). The respective promoter regions were ligated in frame to sites 5' to the GUS gene. A fragment containing bases 2 to 310 from the terminator of the potato proteinase inhibitor (pinII) gene (An et al., *Plant Cell* 1:115-122, 1989) was blunt-end ligated 10 downstream of the GUS coding sequence, to create the GUS expression cassette. The 3' end of the terminator carried a NotI restriction site.

cZ19B1::GUS::pinII was constructed using the above plasmid digested with Ncol, filled in with Klenow enzyme and then digested with NotI to provide insertion sites for the promoter. The plasmid with the isolated cZ19B1 promoter in 15 the TOPOTA cloning vector was digested with Bsal, filled in with Klenow, and digested with NotI. The fragment was ligated into the digested expression cassette and successful subcloning was confirmed by restriction digestion with EcoRI and sequencing.

cim1::GUS::pinII was constructed using the GUS::pinII cassette. The 20 cassette was digested with Ncol, filled in with Klenow and digested again with BamHI. The Cim1 promoter contained in the TOPOTA cloning vector was isolated through digestion with BssHII, filled in with Klenow and subsequently double digested with BamHI/PvuI. The fragment was ligated into the vector and confirmed by restriction analysis and sequencing.

25 mi1ps::GUS::pinII was prepared using the GUS::pinII cassette digested with Ncol, filled in with Klenow, and re-digested with XbaI. The mi1ps promoter in TOPOTA was digested with BsmI, filled in, and then cut with XbaI. The fragment was then ligated into the vector and confirmed by restriction analysis and sequencing.

30 The cim1::GUS::pinII Agrobacterium transformation plasmid was constructed by inserting the GUS expression cassette as a HindIII/NotI fragment and the BAR expression cassette as a NotI/SacI fragment between the right and left T-DNA borders in pSB11 at HindIII and SacI sites. The GUS cassette was inserted proximal to the right T-DNA border. The plasmid pSB11 was obtained

from Japan Tobacco Inc. (Tokyo, Japan). The construction of pSB11 from pSB21 and the construction of pSB21 from starting vectors is described by Komari *et al.* (1996, *Plant J.* 10:165-174). The T-DNA of the plasmid was integrated into the superbinary plasmid pSB1 (Saito *et al.*, EP 672 752 A1) by homologous recombination between the two plasmids. The plasmid pSB1 was also obtained from Japan Tobacco Inc. *E. coli* strain HB101 containing the expression cassettes was mated with Agrobacterium strain LBA4404 harboring pSB1 to create the cointegrate plasmid in Agrobacterium using the method of Ditta *et al.*, (*Proc. Natl. Acad. Sci. USA* 77:7347-7351, 1980). Successful recombination was verified by a Sall restriction digest of the plasmid.

**Example 3: Transformation and Regeneration of Maize Callus via Agrobacterium**  
**Preparation of Agrobacterium suspension:**

Agrobacterium was streaked out from a -80° frozen aliquot onto a plate containing PHI-L medium and cultured at 28°C in the dark for 3 days. PHI-L media comprises 25 ml/l Stock Solution A, 25 ml/l Stock Solution B, 450.9 ml/l Stock Solution C and spectinomycin (Sigma Chemicals) added to a concentration of 50 mg/l in sterile ddH<sub>2</sub>O (stock solution A: K<sub>2</sub>HPO<sub>4</sub> 60.0 g/l, NaH<sub>2</sub>PO<sub>4</sub> 20.0 g/l, adjust pH to 7.0 w/KOH and autoclaved; stock solution B: NH<sub>4</sub>Cl 20.0 g/l, MgSO<sub>4</sub>.7H<sub>2</sub>O 6.0 g/l, KCl 3.0 g/l, CaCl<sub>2</sub> 0.20 g/l, FeSO<sub>4</sub>.7H<sub>2</sub>O 50.0 mg/l, autoclaved; stock solution C: glucose 5.56g/l, agar 16.67 g/l (#A-7049, Sigma Chemicals, St. Louis, MO) and autoclaved).

The plate can be stored at 4°C and used usually for about 1 month. A single colony was picked from the master plate and streaked onto a plate containing PHI-M medium [yeast extract (Difco) 5.0 g/l; peptone (Difco) 10.0 g/l; NaCl 5.0 g/l; agar (Difco) 15.0 g/l; pH 6.8, containing 50 mg/L spectinomycin] and incubated at 28°C in the dark for 2 days.

Five ml of either PHI-A, [CHU(N6) basal salts (Sigma C-1416) 4.0 g/l, Eriksson's vitamin mix (1000X, Sigma-1511) 1.0 ml/l; thiamine.HCl 0.5 mg/l (Sigma); 2,4-dichlorophenoxyacetic acid (2,4-D, Sigma) 1.5 mg/l; L-proline (Sigma) 0.69 g/l; sucrose (Mallinckrodt) 68.5 g/l; glucose (Mallinckrodt) 36.0 g/l; pH 5.2] for the PHI basic medium system, or PHI-I [MS salts (GIBCO BRL) 4.3 g/l; nicotinic acid (Sigma) 0.5 mg/l; pyridoxine.HCl (Sigma) 0.5 mg/l; thiamine.HCl 1.0 mg/l; myo-inositol (Sigma) 0.10 g/l; vitamin assay casamino acids (Difco Lab) 1.0

g/l; 2, 4-D 1.5 mg/l; sucrose 68.50 g/l; glucose 36.0 g/l; adjust pH to 5.2 w/KOH and filter-sterilize] for the PHI combined medium system and 5 ml of 100 mM (3'-5'-Dimethoxy-4'-hydroxyacetophenone, Aldrich chemicals) were added to a 14 ml Falcon tube in a hood. About 3 full loops (5 mm loop size) Agrobacterium was  
5 collected from the plate and suspended in the tube, then the tube vortexed to make an even suspension. One ml of the suspension was transferred to a spectrophotometer tube and the OD of the suspension adjusted to 0.72 at 550 nm by adding either more Agrobacterium or more of the same suspension medium, for an Agrobacterium concentration of approximately  $0.5 \times 10^9$  cfu/ml to  $1 \times 10^9$   
10 cfu/ml. The final Agrobacterium suspension was aliquoted into 2 ml microcentrifuge tubes, each containing 1 ml of the suspension. The suspensions were then used as soon as possible.

Embryo isolation, infection and co-cultivation:

15 About 2 ml of the same medium (here PHI-A or PHI-I) used for the Agrobacterium suspension were added into a 2 ml microcentrifuge tube. Immature embryos were isolated from a sterilized ear with a sterile spatula (Baxter Scientific Products S1565) and dropped directly into the medium in the tube. A total of about 100 embryos were placed in the tube. The optimal size of the  
20 embryos was about 1.0-1.2 mm. The cap was then closed on the tube and the tube vortexed with a Vortex Mixer (Baxter Scientific Products S8223-1) for 5 sec. at maximum speed. The medium was removed and 2 ml of fresh medium were added and the vortexing repeated. All of the medium was drawn off and 1 ml of Agrobacterium suspension added to the embryos and the tube vortexed for 30  
25 sec. The tube was allowed to stand for 5 min. in the hood. The suspension of Agrobacterium and embryos was poured into a Petri plate containing either PHI-B medium [CHU(N6) basal salts (Sigma C-1416) 4.0 g/l; Eriksson's vitamin mix (1000X, Sigma-1511) 1.0 ml/l; thiamine.HCl 0.5 mg/l; 2,4-D 1.5 mg/l; L-proline 0.69 g/l; silver nitrate 0.85 mg/l; gelrite (Sigma) 3.0 g/l; sucrose 30.0 g/l;  
30 acetosyringone 100 mM; pH 5.8], for the PHI basic medium system, or PHI-J medium [MS Salts 4.3 g/l; nicotinic acid 0.50 mg/l; pyridoxine HCl 0.50 mg/l; thiamine.HCl 1.0 mg/l; myo-inositol 100.0 mg/l; 2, 4-D 1.5 mg/l; sucrose 20.0 g/l; glucose 10.0 g/l; L-proline 0.70 g/l; MES (Sigma) 0.50 g/l; 8.0 g/l agar (Sigma A-7049, purified) and 100 mM acetosyringone with a final pH of 5.8 for the PHI

combined medium system. Any embryos left in the tube were transferred to the plate using a sterile spatula. The Agrobacterium suspension was drawn off and the embryos placed axis side down on the media. The plate was sealed with Parafilm tape or Pylon Vegetative Combine Tape (product named "E.G.CUT" and 5 is available in 18 mm x 50 m sections; Kyowa Ltd., Japan) and incubated in the dark at 23-25°C for about 3 days of co-cultivation.

Resting, selection and regeneration steps:

For the resting step, all of the embryos were transferred to a new plate 10 containing PHI-C medium [CHU(N6) basal salts (Sigma C-1416) 4.0 g/l; Eriksson's vitamin mix (1000X Sigma-1511) 1.0 ml/l; thiamine.HCl 0.5 mg/l; 2.4-D 1.5 mg/l; L-proline 0.69 g/l; sucrose 30.0 g/l; MES buffer (Sigma) 0.5 g/l; agar (Sigma A-7049, purified) 8.0 g/l; silver nitrate 0.85 mg/l; carbenicillin 100 mg/l; pH 5.8]. The plate was sealed with Parafilm or Pylon tape and incubated in the dark at 28°C for 15 3-5 days.

Longer co-cultivation periods may compensate for the absence of a resting step since the resting step, like the co-cultivation step, provides a period of time for the embryo to be cultured in the absence of a selective agent. Those of ordinary skill in the art can readily test combinations of co-cultivation and resting 20 times to optimize or improve the transformation

For selection, all of the embryos were then transferred from the PHI-C medium to new plates containing PHI-D medium, as a selection medium, [CHU(N6) basal salts (SIGMA C-1416) 4.0 g/l; Eriksson's vitamin mix (1000X, Sigma-1511) 1.0 ml/l; thiamine.HCl 0.5 mg/l; 2.4-D 1.5 mg/l; L-proline 0.69 g/l; sucrose 30.0 g/l; MES buffer 0.5 g/l; agar (Sigma A-7049, purified) 8.0 g/l; silver 25 nitrate 0.85 mg/l; carbenicillin (ICN, Costa Mesa, CA) 100 mg/l; bialaphos (Meiji Seika K.K., Tokyo, Japan) 1.5 mg/l for the first two weeks followed by 3 mg/l for the remainder of the time.; pH 5.8] putting about 20 embryos onto each plate. The plates were sealed as described above and incubated in the dark at 28°C for the 30 first two weeks of selection. The embryos were transferred to fresh selection medium at two-week intervals. The tissue was subcultured by transferring to fresh selection medium for a total of about 2 months. The herbicide-resistant calli were then "bulked up" by growing on the same medium for another two weeks until the diameter of the calli was about 1.5-2 cm.

For regeneration, the calli were then cultured on PHI-E medium [MS salts 4.3 g/l; myo-inositol 0.1 g/l; nicotinic acid 0.5 mg/l, thiamine.HCl 0.1 mg/l, Pyridoxine.HCl 0.5 mg/l, Glycine 2.0 mg/l, Zeatin 0.5 mg/l, sucrose 60.0 g/l, Agar (Sigma, A-7049) 8.0 g/l, Indoleacetic acid (IAA, Sigma) 1.0 mg/l, Abscisic acid (ABA, Sigma) 0.1 mM, Bialaphos 3 mg/l, carbenicillin 100 mg/l adjusted to pH 5.6] in the dark at 28°C for 1-3 weeks to allow somatic embryos to mature. The calli were then cultured on PHI-F medium (MS salts 4.3 g/l; myo-inositol 0.1 g/l; Thiamine.HCl 0.1 mg/l, Pyridoxine.HCl 0.5 mg/l, Glycine 2.0 mg/l, nicotinic acid 0.5 mg/l; sucrose 40.0 g/l; gelrite 1.5 g/l; pH 5.6] at 25°C under a daylight schedule of 16 hrs. light (270 uE m<sup>-2</sup>sec<sup>-1</sup>) and 8 hrs. dark until shoots and roots developed. Each small plantlet was then transferred to a 25x150 mm tube containing PHI-F medium and grown under the same conditions for approximately another week. The plants were transplanted to pots with soil mixture in a greenhouse. GUS+ events are determined at the callus stage or regenerated plant stage.

For Hi-II a preferred optimized protocol was 0.5 x 10<sup>9</sup> cfu/ml Agrobacterium, a 3-5 day resting step, and no AgNO<sub>3</sub> in the infection medium (PHI-A medium).

**Example 4:** In situ Localization of Cim1 mRNA in 5 DAP Maize Kernel.

In situ hybridization was performed using the protocol of Jackson, D.P. (1991) *In situ Hybridization in Plants, Molecular Plant Pathology: A Practical Approach*, D.J. Bowles, S. J. Gurr, and M. McPherson, eds. Oxford University Press, England, pp.63-74. Both a sense and antisense probe corresponding to a protein of the Cim1 cDNA were used. Probes were labelled non-isotopically with digoxigenin and incubated with various sections of 5 DAP (days after pollination) maize kernels which had been fixed and embedded. Following extensive washing to remove unbound probe, sections were incubated with anti-digoxigenin alkaline phosphatase to detect areas of probe hybridization. For Cim1, mRNA was detected specifically with the antisense probe and restricted to nucellus tissue. The sense control probe did not hybridize.

**Example 5:** Northern Analysis of Gene Expression in Vegetative Tissue and Developing Kernels.

Total RNA (10 g) was size fractionated on a 1% formaldehydeagarose gel and transferred to a nitrocellulose membrane. Membranes were hybridized under stringent conditions with 32P-labelled probes representing cDNA fragments of the various genes. After extensive washing to remove unbound probe, membranes 5 were exposed on X-ray film. RNA samples were obtained from vegetative tissues as well as developing maize kernels.

The Cim1 expression pattern showed no expression in vegetative tissues or in the isolated embryo or endosperm of developing kernels. Cim1 was predominantly expressed in the early (5 DAP) whole kernel. The gene for mi1ps 10 was expressed predominately in the embryo of developing kernels ( 13-40 DAP) while cZ19B1 was expressed in the endosperm of mid-late kernel development (13-40 DAP). Neither mi1ps nor cZ19B1 were expressed in vegetative tissues.

All publications and patent applications mentioned in the specification are indicative of the level of those skilled in the art to which this invention pertains. All 15 publications and patent applications are herein incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be 20 obvious that certain changes and modifications may be practiced within the scope of the appended claims.

## WHAT IS CLAIMED:

1. An isolated promoter that is capable of driving transcription in a seed-preferred manner, wherein said promoter comprises a nucleotide sequence selected from the group consisting of:
  - a) sequences capable of driving expression of coding regions selected from the group consisting of coding regions for Cim 1 (cytokinin-induced message); cZ19B1 (maize 19KDa zein); or mi1ps (myo-inositol-1-phosphate synthase);
  - b) a sequence comprising at least 20 contiguous nucleotides of the sequence set forth in either of SEQ ID NOS 1, 4, or 7;
  - c) a sequence comprising a variant or fragment of the nucleotide sequence set forth in either of SEQ ID NOS 1, 4, or 7;
  - d) the nucleotide sequences set forth in SEQ ID NOS 1, 4, or 7;
  - e) nucleic acids having at least 60% sequence identity to SEQ ID NO 1, at least 81% sequence identity to SEQ ID NO 4, or at least 60% sequence identity to SEQ ID NO 7, wherein the % sequence identity is based on the entire sequence and is determined by GAP version 10 analysis using default parameters;
  - f) a sequence that hybridizes to any one of SEQ ID NOS: 1, 4, or 7, under stringent conditions; and
  - g) a nucleic acid complementary to a nucleic acid of (a) through (f).
2. An expression cassette comprising a promoter and a nucleotide sequence operably linked to said promoter, wherein said promoter is capable of initiating seed-preferred transcription of said nucleotide sequence in a plant cell, wherein said promoter comprises a nucleotide sequence selected from the group consisting of:
  - a) sequences capable of driving expression of coding regions selected from the group consisting of coding regions for Cim 1 (cytokinin-induced message); cZ19B1 (maize 19KDa zein); or mi1ps (myo-inositol-1-phosphate synthase);

- b) a sequence comprising at least 20 contiguous nucleotides of the sequence set forth in either of SEQ ID NOS 1, 4, or 7;
- c) a sequence comprising a variant or fragment of the nucleotide sequence set forth in either of SEQ ID NOS 1, 4, or 7;
- 5 d) the nucleotide sequences set forth in SEQ ID NOS 1, 4, or 7;
- e) nucleic acids having at least 60% sequence identity to SEQ ID NO 1, at least 81% sequence identity to SEQ ID NO 4, or at least 60% sequence identity to SEQ ID NO 7, wherein the % sequence identity is based on the entire sequence and is determined by GAP version 10 analysis using default parameters;
- 10 f) a sequence that hybridizes to any one of SEQ ID NOS: 1, 4, or 7, under stringent conditions; wherein expression of said sequence changes the phenotype of seed of a plant transformed with said expression cassette; and
- 15 g) a nucleic acid complementary to a nucleic acid of (a) through (f).
3. A transformation vector comprising an expression cassette, said expression cassette comprising a maize promoter and a nucleotide sequence operably linked 20 to said promoter, wherein said promoter is capable of initiating seed-preferred transcription of said nucleotide sequence in a plant cell, wherein said promoter comprises a nucleotide sequence selected from the group consisting of:
- a) sequences capable of driving expression of coding regions selected from the group consisting of coding regions for Cim 1 (cytokinin-induced message); cZ19B1 (maize 19KDa zein); or mi1ps (myo-inositol-1-phosphate synthase);
- 25 b) a sequence comprising at least 20 contiguous nucleotides of the sequence set forth in either of SEQ ID NOS 1, 4, or 7;
- c) a sequence comprising a variant or fragment of the nucleotide sequence set forth in either of SEQ ID NOS 1, 4, or 7;
- 30 d) the nucleotide sequences set forth in SEQ ID NOS 1, 4, or 7;
- e) nucleic acids having at least 60% sequence identity to SEQ ID NO 1, at least 81% sequence identity to SEQ ID NO 4, or at least 60% sequence identity to SEQ ID NO 7, wherein the % sequence identity is based on the entire sequence and is determined by GAP version 10

analysis using default parameters;

- f) a sequence that hybridizes to any one of SEQ ID NOS: 1, 4, or 7, under stringent conditions; and
- 5 g) a nucleic acid complementary to a nucleic acid of (a) through (f).

4. A plant stably transformed with an expression cassette comprising a maize promoter and a nucleotide sequence operably linked to said promoter, wherein said promoter is capable of initiating seed-preferred transcription of said 10 nucleotide sequence in a plant cell, wherein said promoter comprises a nucleotide sequence selected from the group consisting of:
  - a) sequences capable of driving expression of coding regions selected from the group consisting of coding regions for Cim 1 (cytokinin-induced message); cZ19B1 (maize 19KDa zein); or mi1ps (myo-inositol-1-phosphate synthase);
  - 15 b) a sequence comprising at least 20 contiguous nucleotides of the sequence set forth in either of SEQ ID NOS 1, 4, or 7;
  - c) a sequence comprising a variant or fragment of the nucleotide sequence set forth in either of SEQ ID NOS 1, 4, or 7;
  - d) the nucleotide sequences set forth in SEQ ID NOS 1, 4, or 7;
  - 20 e) nucleic acids having at least 60% sequence identity to SEQ ID NO 1, at least 81% sequence identity to SEQ ID NO 4, or at least 60% sequence identity to SEQ ID NO 7, wherein the % sequence identity is based on the entire sequence and is determined by GAP version 10
  - 25 analysis using default parameters;
    - f) a sequence that hybridizes to any one of SEQ ID NOS: 1, 4, or 7, under stringent conditions; wherein expression of the nucleotide sequences alters the phenotype of seed of said transformed plant; and
    - 30 g) a nucleic acid complementary to a nucleic acid of (a) through (f).
5. The plant of claim 4, wherein said plant is a monocot.
6. The plant of claim 5, wherein said monocot is corn, wheat, rice, barley, 35 sorghum, or rye.

7. The plant of claim 4, wherein said plant is a dicot.

8. Seed of the plants of claim 4.

5 9. A method for selectively expressing a nucleotide sequence in a plant seed, said method comprising transforming a plant cell with a transformation vector comprising an expression cassette, said expression cassette comprising a maize promoter and a nucleotide sequence operably linked to said promoter, wherein said promoter is capable of initiating seed-preferred transcription of said  
10 nucleotide sequence in a plant cell, wherein said promoter comprises a nucleotide sequence selected from the group consisting of:

- a) sequences capable of driving expression of coding regions selected from the group consisting of coding regions for Cim 1 (cytokinin-induced message); cZ19B1 (maize 19KDa zein); or mi1ps (myo-inositol-1-phosphate synthase);
- b) a sequence comprising at least 20 contiguous nucleotides of the sequence set forth in either of SEQ ID NOS 1, 4, or 7;
- c) a sequence comprising a variant or fragment of the nucleotide sequence set forth in either of SEQ ID NOS 1, 4, or 7;
- 20 d) the nucleotide sequences set forth in SEQ ID NOS 1, 4, or 7;
- e) nucleic acids having at least 60% sequence identity to SEQ ID NO 1, at least 81% sequence identity to SEQ ID NO 4, or at least 60% sequence identity to SEQ ID NO 7, wherein the % sequence identity is based on the entire sequence and is determined by GAP version 10 analysis using default parameters;
- f) a sequence that hybridizes to any one of SEQ ID NOS: 1, 4, or 7, under stringent conditions; and
- 30 g) a nucleic acid complementary to a nucleic acid of (a) through (f).

10. The method of claim 9 further comprising regenerating a stably transformed plant from said transformed plant cell; wherein expression of said nucleotide sequences alters the phenotype of said plant seed.

11. The method of claim 9, wherein said heterologous nucleotide sequence encodes a gene involved in fatty acid synthesis.

12. The method of claim 9, wherein said heterologous nucleotide sequence 5 encodes a gene having enhanced amino acid content.

13. A plant cell stably transformed with an expression cassette comprising a maize promoter and a nucleotide sequence operably linked to said promoter, wherein said promoter is capable of initiating seed-preferred transcription of said 10 nucleotide sequence in a plant cell, wherein said promoter comprises a nucleotide sequence selected from the group consisting of:

- a) sequences capable of driving expression of coding regions selected from the group consisting of coding regions for Cim 1 (cytokinin-induced message); cZ19B1 (maize 19KDa zein); or mi1ps (myo-inositol-1-phosphate synthase);
- b) a sequence comprising at least 20 contiguous nucleotides of the sequence set forth in either of SEQ ID NOS 1, 4, or 7;
- c) a sequence comprising a variant or fragment of the nucleotide sequence set forth in either of SEQ ID NOS 1, 4, or 7;
- d) the nucleotide sequences set forth in SEQ ID NOS 1, 4, or 7;
- e) nucleic acids having at least 60% sequence identity to SEQ ID NO 1, at least 81% sequence identity to SEQ ID NO 4, or at least 60% sequence identity to SEQ ID NO 7, wherein the % sequence identity is based on the entire sequence and is determined by GAP version 10 analysis using default parameters;
- f) a sequence that hybridizes to any one of SEQ ID NOS: 1, 4, or 7, under stringent conditions; wherein expression of the nucleotide sequences alters the phenotype of seed of said transformed plant cell; and
- g) a nucleic acid complementary to a nucleic acid of (a) through 30 (f).

14. The plant cell of claim 13, wherein said plant cell is from a monocotyledonous plant.

15. The plant cell of claim 14, wherein said plant is corn, wheat, rice, barley, sorghum, or rye.
16. The plant cell of claim 13, wherein said plant cell is from a dicotyledonous plant.

## SEQUENCE LISTING

<110> Pioneer Hi-Bred International, Inc.

<120> Seed-Preferred Promoters

<130> 0869-PCT

<150> US 60/097,233

<151> 1998-08-20

<160> 9

<170> FastSEQ for Windows Version 3.0

<210> 1

<211> 1609

<212> DNA

<213> Zea mays

<220>

<221> promoter

<222> (0)...(1609)

<400> 1

actggcggca ctgaaaaacg ccagtgcaaa tagatctaga accgtacta tagagttct	60
gtgtactagt atagtcgggt gagccaaaca gtcatgtgt ccctagtgca tcacaacagg	120
tgtcttctgt tgcgctagcg gacaagttag agatcgatta taaagatgtctccagctata	180
tattgtatca tggcgacgtc tacactgggt ctgatcgctg cggttatgca gtggtgatct	240
gtctctctct tccattttga ggcgcgaggt tgggttttt ttccctgttg gggatcgga	300
ttcgaggatg atgaaacgta acggccagca agctagttgt agcatttctg caccactaca	360
catgcattgtt cgatctgtgt atatatatac cggcatacac agatgtgtgt atataccggc	420
atacacagag gcatatatat gcatatgtca atctcggtca tgtacaccct gtagaattct	480
gctgtcttcg tccgatcaca catggcgccg gtcatcgccg tattctgtca gctgcaaata	540
caatacgcta caagatgggc gtcacatgca tatacatgtta actaggccac gcgtgtgtgc	600
cacgcattgtt gatgctgaca tgcattgacat taatgaacaa ttgtaggatg gcttcaactg	660
cgaggttctt gcatcatcgat ggttttacaa ttttcaataa acacggacat acacacacac	720
tgaactatttta gattgtttttt accatttcgat tattgttatttccatcatctg agacttttat	780
atatttgcgc acttacatgtt gctacaaatg ttaccataaa tataaaccat gcatcagact	840
tgcacaaaaa ttaccaagga tagagagtgg ggataaaaag gggatctga acagttacga	900
taaattttat atttcaaaat atagatgtat gaatttttagt gctaatcttt tcttacatttgc	960
ttatgaagca cattgttattt cgacaggatt aaatttttgc ataaactgtttt ttacattatg	1020
tactccaaaaaa aaagtggaaaa tatctgtata ttggatattt atatctatca tctaagaaga	1080
tatacgatag atttgaagtt tacattttat gaatcttacaa aagctcaatg ctaaaaaaaca	1140
agaatatgaa ttgcgcagc caattctata tccttatttaa ttatttcacag tcaaaggaaaa	1200
aaagagaaca aagttaacagg tacgcggaga agtattttttt ttcattcccta accatttagga	1260
tggatttgtt agagaaatattt ttagggacgtt ataaaggaaat taaaataaaaa agattaagaa	1320
cctaacatgt gtaagttaaag gaaatatttattt catatagaag aaaaggaaaa tccaaatcagc	1380
tgtggaaagaa ttaatttagaa accgggattttt atgaaactcg agcaacgtca atataatgtt	1440
gtccttctgtt aaatttcagt gtcacggaaag gtgcacgggc cggtataaaa tagatgccat	1500
gacaccatgc accacaccccc gcatctccta gccccagcta gtagatcattt ggacacatata	1560
cgaaacagct ctctagctag cgccgtcgcc gcttaacctg gcttccggc	1609

<210> 2

<211> 26

<212> DNA

<213> Zea mays

<pre> &lt;400&gt; 2 ggtgatggca atgaagcagg tcaaga </pre>	26
<pre> &lt;210&gt; 3 &lt;211&gt; 28 &lt;212&gt; DNA &lt;213&gt; Zea mays </pre>	
<pre> &lt;400&gt; 3 tggaaagccat gccggaagcc aggttaga </pre>	
<pre> &lt;210&gt; 4 &lt;211&gt; 921 &lt;212&gt; DNA &lt;213&gt; Zea mays </pre>	
<pre> &lt;220&gt; &lt;221&gt; promoter &lt;222&gt; (1)...(922) </pre>	
<pre> &lt;400&gt; 4 gcttggtcgc chtagcagca ctgcagaaa gaccaaggaa gcataaggag gcaaaatatt ttggctgcca ttgttgtac actattgtgc tttctatgtt gctagtttgt ctcaataata tgtgcgatat attagctact aatttaagag atatttatac acatctcagt ccttgtgctt gtgcaagggtt aattttggat ttgcaaaaaa gaatattagt aagataaggt agagatgaca tgttatctag ttggtttga catttagtga ttactttgtt gggttataat ttatacgtaa gcaagacaaa cttgataaaa tacataataat aagtggttt tcctctcggt attgtccaca tacatgggtg acaatggtgt gatgcaatac ctttacacat gtgtAACGAT agcccacata tgtgttgctt ggtctggcta ttttccacta gaaaaaaaaat catggcacca ttccaccattg ttaatttttag tgtggtaac tttgtgttaa taatcggtt aaagaatcac tcgatcccc ccgatagaaaa ttttaggtact ttttagttca ttttaatcta taaagttgat actgatttt ttagtgcac ataaaggcgc ctatgtgatt taaacataat gaatatgagc aaaataaata gtaacatata aggaagacat tgctagttga tatagtcgta ctggttgtaa tagtgcagtg ttgttgctat gtcgactcaa tcacaacaaa aaacactagg aagaagcggg ggaagtaagg atgtgtgtcg ctgatatttc octagaacac gattgtcgtc ctcatctga gtagatattc atttcggagg cctgctcagt gctgtaatgg agtaccagcc cggccgtcg accacgcgtg ccctatagtg agtcgtatta c </pre>	
<pre> &lt;210&gt; 5 &lt;211&gt; 30 &lt;212&gt; DNA &lt;213&gt; Zea mays </pre>	
<pre> &lt;400&gt; 5 atattttgcc ttccttatgc tcccttggtc </pre>	
<pre> &lt;210&gt; 6 &lt;211&gt; 31 &lt;212&gt; DNA &lt;213&gt; Zea mays </pre>	
<pre> &lt;400&gt; 6 tttctgcaag tgctgctacg gcgaccaagc c </pre>	
<pre> &lt;210&gt; 7 &lt;211&gt; 991 &lt;212&gt; DNA &lt;213&gt; Zea mays </pre>	

<220>  
 <221> promoter  
 <222> (1)...(991)

<400> 7

gaattgggcc	ctctagatgc	atgctcgagc	ggccgccagt	gtgatggata	tctgcagaat	60
tcggcttcta	atacgactca	ctataggca	cgcgtgtcg	atggcccggc	tggtaacttt	120
gaaacaaaaa	aataagaaga	ataagagata	atttggaaac	tctatTTTT	tcgagggatt	180
tttatttttc	taaaaaagta	actaaatTTT	attagaaaaa	taaaaatcat	ttaagaaaaat	240
gaggtaatg	gaaaccatt	tttccaaaag	atTTTTATT	ttctaagaaaa	aatttagTTA	300
ttttccttcta	tagggaaaca	gaaatacccc	aaatAGTTC	taaaactAGC	actatttgcA	360
aactcaaact	taatgggtgt	ttggTTTcta	gggactaatG	tttaagtccc	ttcattttat	420
ttcttttag	tatataaaatt	gctaaatata	gaaactaaaa	taaagtTTTA	gtttttatAT	480
ttacaatTTT	agaactaaaa	tataataaaa	tgtagggact	aaacattAGT	ccctacAAAC	540
caaacacGCC	cttgtaattt	tccatcggtA	tgtcCTCCGT	cttcccgaaa	gccttcattc	600
ggctcgctgt	cttcttctA	tcgctcacAA	ccaacaacca	tagccccacAG	gcccGCCGGCG	660
gccGCCAGCC	gacgtcctcc	catttcgctc	ccccctcctcc	gctgcgggtcg	agcaaaAGTT	720
ccggccatcc	ggcaatcccc	cgccacCCGG	cggttcaaat	cgtatcttc	tgacctgacG	780
cggctacgac	gtcgcttc	cgtcccttc	gatccggTgg	ggtccgTTTC	tttcaaggcG	840
ggcctcgctg	gccgcctcgT	ggcagtgacc	gtcgaaccct	ctataaatcc	cgtgccccGA	900
gcacccttcc	tcgatcacac	aacccaaggc	agccacagca	gcctcTTCC	tcctctcaCT	960
ctcgctcgTg	ctcgctcgC	tacctcgctt	C			991

<210> 8  
 <211> 26  
 <212> DNA  
 <213> Zea mays

<400> 8

ctctcgacgc	ggaagctctc	gatgaa	26
------------	------------	--------	----

<210> 9  
 <211> 26  
 <212> DNA  
 <213> Zea mays

<400> 9

catcttgcct	ttcctccccc	cttttc	26
------------	------------	--------	----

## INTERNATIONAL SEARCH REPORT

International Application No  
PCT/US 99/18628

A. CLASSIFICATION OF SUBJECT MATTER  
IPC 7 C12N15/29 C12N15/61 C12N15/82 C12N5/10 C07K14/415  
A01H5/00

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)  
IPC 7 C12N A01H C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the International search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 91 13993 A (UPJOHN CO) 19 September 1991 (1991-09-19) see esp. examples 5-11	1-16
X	WO 98 26064 A (DEKALB GENETICS CORP) 18 June 1998 (1998-06-18) the whole document	1-16
X	KRIZ A. ET AL.: "AC X05911" EMBL DATABASE, 2 April 1988 (1988-04-02), XP002126294 HEIDELBERG the whole document	1-16 -/-

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

## \* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the International filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the International filing date but later than the priority date claimed

- "T" later document published after the International filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "8" document member of the same patent family

Date of the actual completion of the International search

20 December 1999

Date of mailing of the International search report

11/01/2000

Name and mailing address of the ISA  
European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2220 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 51 051 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Kania, T

## INTERNATIONAL SEARCH REPORT

Intr. Int'l Application No.  
PCT/US 99/18628

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	GIOVINAZZO G. ET AL.: "Functional analysis of the regulatory region of a zein gene in transiently transformed protoplasts" PLANT MOLECULAR BIOLOGY, vol. 19, 1992, pages 257-263, XP002126295 the whole document _____	1-16
A	US 5 677 474 A (ROGERS JOHN C) 14 October 1997 (1997-10-14) the whole document _____	1-16
A	CROWELL D.: "Cytokinin regulation of a soybean pollen allergen gene" PLANT MOLECULAR BIOLOGY, vol. 25, 1994, pages 829-835, XP002126296 the whole document _____	1-16
A	WO 98 30698 A (CALGENE INC ;PEAR JULIE R (US); STALKER DAVID M (US)) 16 July 1998 (1998-07-16) the whole document _____	1-16
T	LARSON S. AND RABOY V.: "Linkage mapping of maize and barley myo-inositol 1-phosphate synthase DNA sequences: correspondence with low phytic acid mutation" THEORETICAL AND APPLIED GENETICS, vol. 99, July 1999 (1999-07), pages 27-36, XP000863631 see esp. pp.32/33 _____	1-16

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 99/18628

### Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1.  Claims No.: because they relate to subject matter not required to be searched by this Authority, namely:
  
2.  Claims No.: because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
  
3.  Claims No.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

### Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this International application, as follows:

see additional sheet

1.  As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
  
2.  As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
  
3.  As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims No.:
  
4.  No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims No.:

Remark on Protest

- The additional search fees were accompanied by the applicant's protest.
- No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows :

1. Claims: 1-16 partially

An isolated promoter, esp. from maize, capable of driving transcription in a seed-preferred manner selected from: sequences capable of driving expression of the coding region for Cim1, sequences comprising at least 20 contiguous nucleotides of SEQ ID NO:1, or variants, fragments, homologs thereof, and hybridizing and complementary sequences. Expression cassettes and transformation vectors comprising said promoters, as well as plant cells and plants stably transformed therewith.

Methods for the selective expression of a nucleotide in a plant seed employing said promoters, expression cassettes, or vectors, especially driving genes involved in fatty acid synthesis, or encoding proteins with enhanced amino acids content.

2. Claims: 1-16 partially

idem for cZ19B1 (maize 19KDa zein), SEQ ID NO:4

3. Claims: 1-16 partially

idem for milps (myo-inositol-1-phosphate synthase), SEQ ID NO:7

# INTERNATIONAL SEARCH REPORT

Information on patent family members

Int'l	Serial Application No
	PCT/US 99/18628

Patent document cited in search report	Publication date	Patent family member(s)			Publication date
WO 9113993	A 19-09-1991	AU 7583691 A	EP 0519011 A		10-10-1991 23-12-1992
WO 9826064	A 18-06-1998	AU 5795898 A	EP 0942971 A		03-07-1998 22-09-1999
US 5677474	A 14-10-1997	AU 638409 B	AU 4037289 A	EP 0428572 A	01-07-1993 05-03-1990 29-05-1991
		JP 4500153 T	WO 9001551 A		16-01-1992 22-02-1990
WO 9830698	A 16-07-1998	AU 5732298 A			03-08-1998